DOCUMENT RESUME

ED 453 220 TM 032 571

AUTHOR Lumley, Dale R.; Yan, Wenfan
TITLE The Impact of State Mandated, Large-Scale Writing Assessment Policies in Pennsylvania.
PUB DATE 2001-04-11
NOTE 40p.; Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 10-14, 2001).
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Educational Policy; Educational Practices; Elementary Secondary Education; Holistic Approach; Scoring; State Programs; Teacher Attitudes; *Teachers; *Teaching Methods; *Test Use; Testing Programs; *Writing Instruction; Writing Tests
IDENTIFIERS *Pennsylvania

ABSTRACT

This study examines the impact of the Pennsylvania writing assessment policy on writing instruction and teaching methodology. Drawing on self-reported teacher perceptions and classroom practices, the study also attempts to identify the factors that influence teacher beliefs and their use of materials and processes related to the state writing assessment. The field-tested survey, which contained 60 items, was completed by 168 teachers from 20 schools, 56.9% of whom taught elementary school. The initial study examines the effects of 16 factors on classroom practices and teaching methodology, but this analysis focuses on three key factors: teaching responsibility, holistic training, and Pennsylvania System of School Assessment scores at or above the band of similar schools. Teacher responses to the survey indicate that they are providing students with frequent writing experiences and opportunities. The state assessment seems to have motivated teachers to present these experiences. Teachers also appear to recognize the value of holistic scoring and the characteristics of effective writing as presented in the Pennsylvania Holistic Scoring Guide. However, it was evident that even though many teachers agreed with the beliefs and values of holistic scoring and the characteristics of effective writing, they were reluctant to use the state rubric, descriptors, and writing samples as the basis for classroom instruction. This finding may suggest: a weakness exists in the supporting materials provided by the state, teachers are developing their own rubrics and samples, or that teachers are clinging to traditional methods and not adopting a progressive approach to writing instruction. (Contains 4 tables and 39 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 453 220

THE IMPACT OF STATE MANDATED, LARGE-SCALE WRITING ASSESSMENT POLICIES IN PENNSYLVANIA

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
D. Lumley
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

Dale R. Lumley and Wenfan Yan
Indiana University of Pennsylvania

Paper Presented at the American Educational Research Association Annual Conference
Seattle, WA – April 11, 2001

TM032571

BEST COPY AVAILABLE

The Impact of State Mandated, Large-Scale Writing Assessment Policies in Pennsylvania

Dale R. Lumley and Wenfan Yan
Indiana University of Pennsylvania

Introduction

There has been a growing trend for states to adopt standards and to implement state exams as a way to assess school achievement. While the quality of state standards continues to be debated (Finn & Petrilli 2000), public schools must come to grips with the policies and procedures of state exams and the ramifications of scores as a reflection of school quality. High stakes testing and issues of accountability continue to be debated (AERA Policy Statement, 2000; Gordon, 2000; Koretz & Barron, 1998), yet little has been determined regarding the influence which state mandated assessment policies have on actual classroom practices. A growing number of studies have examined this issue (Firestone, Mayrowetz & Fairman, 1998; Fuchs et al., 1999; Torrance, 1993). Relatively few studies (Bridge et al., 1997; Bridge & Hiebert, 1985), however, have examined the influence that large-scale writing assessments have had on instruction.

The topic of writing instruction and appropriate assessment remains an important issue for teachers, administrators, researchers and policy makers. Relevant studies (Atwell, 1987; Barnes, 1975; Berthoff, 1982; Britton, 1970; Graves, 1983; Rief, 1992; Ruth & Murphy, 1988; and Tchudi, 1991) have shown a growing understanding for the sophisticated ways by which students learn to write and the great divergence of methods undertaken by practitioners in the field. Previous studies have suggested that large-scale writing assessments, which employ some type of holistic scoring based on pre-established criteria and which extend beyond the classroom to include entire buildings, districts or states, provide an accurate and valid measure of student writing (Farr & Jongsma, 1993; Novak et al., 1996; Soltis & Walberg, 1989; Walberg &

Ethington, 1991). What has been missing from those studies, however, is the impact that such large-scale assessment has had on instructional practices and teaching strategies.

While the Pennsylvania Department of Education mandated large-scale writing assessments in 1990, little has been examined on how those assessments have influenced instruction. For the past decade, the state emphasized holistic scoring as a way to gauge district programs and school effectiveness regarding writing skills and instruction. More recently, however, the writing assessment has moved to domain scoring as a way to provide individual writers with feedback as well as to determine incentive funding. The state mandated writing assessment has been challenged in the courts and several school districts have united in an effort to halt the writing assessment. One reason for these challenges and the uncertainty regarding large-scale writing assessments and their place in the public school is the lack of a thorough review of how such assessments have altered classroom practices and student writing skills. This study examines the impact of the Pennsylvania writing assessment policy on writing instruction and teaching methodology. Based on teacher, self-reported perceptions and classroom practices, the study also attempts to identify the factors, which influence teacher beliefs, and their use of materials and processes related to the state writing assessment.

The Assessment-Instruction Connection

Evident from a broad body of literature related to writing and instruction is that the metacognitive processes of students are important to their growth and development as writers and learners. In addition, these studies indicate that teachers must reflect upon assessment results in order to shape, alter and improve classroom practices. Both writing and teaching can be viewed as non-linear processes in which reflection and self-assessment are key components. The better

one integrates the assessment with desired skills and outcomes, the more successful the performance, whether that performance is a student composition or a teacher lesson, method or strategy (Stiggins, 1997; Wiggins, 1993). These concepts are keenly important to the teacher of writing. "Assumptions about the nature of writing and models of the writing process have consequences for the design of assessment instruments. If writing is conceived of as a loose aggregation of skills, we may be satisfied with discrete point, multiple choice measures; however, if writing is conceived of as a complex, purposeful, cognitive act, we must demand writing tasks that draw on the communicative and cognitive processes" (Ruth & Murphy, 1988, p. 110.) More dramatically, cognitivist Joseph Petraglia (1995) believes that only when writing is viewed as a particularly challenging ill-structured problem can a cognitive framework be applied. Furthermore, Petraglia believes that most general writing instruction in a school setting ignores the cognitive perspective. The potential exists, however, for large-scale assessments which recognize the complex cognitive nature of writing, and which are gaining political support, to be a driving force of change in classroom writing instruction.

Paris and Winograd (1990) see the need for metacognition to provide "cognitive tools for accomplishing the craft of schooling" (p. 90). From their perspective, students and teachers can be taught particular strategies to accomplish specific purposes. The teacher's role is to inform students about effective problem-solving strategies by making them aware of cognitive and motivational characteristics of thinking and writing (Paris & Winograd, 1990). That same premise drives this study. Only if teachers reflect upon assessment results and work to redefine personal and educational philosophies while seeking new ways to implement and improve classroom practices will large-scale assessments have much bearing toward improving instruction or increasing learning.

Current research on assessment suggests that when teachers and students are an integral part of on-going assessment strategies they become more knowledgeable and skilled which in turn allows them to perform at higher levels of proficiency (McTighe, 1997; Stiggins, 1997; Wiggins, 1993). If one truly knows the skills that are assessed and the criteria by which such performances are judged, then one can work to improve those skills in order to achieve greater performance. Vygotsky (1962) feels such assessment practices should challenge students and teachers to measure their progress toward explicit goals. If necessary, they can reallocate their time to different activities in order to reach those goals. But have the changes in writing assessment, especially large-scale holistic scoring, precipitated any change in classroom instruction? Have teachers, indeed, reallocated their time or altered their practices in light of performance assessment results?

Many studies indicate that writing instruction as part of academic courses runs contrary to the goals of large-scale assessment. Ruth and Murphy (1988) point out that little of the writing instruction provided by teachers helps to prepare students for the kinds of writing tasks they face in large-scale assessments. Martin (1992) recognized that the energy and imagination needed to promote language across the curriculum and to bring about real change has been "sucked away into the insatiable maw of the national curriculum" (p. 20). Will the same fate await new directions in assessment? Or will the state and national support for large-scale writing assessment continue to be refined to shape the next generation of language and writing acquisition research?

Edward White (1994) writes, "as long as assessment remains a peripheral concern for us, an afterthought to our curriculum, we will neglect the opportunities for test development that we so richly possess" (p.66) He goes on to explain that much of the writing required in school deals with retelling something the teacher said or summarizing something students read. The emphasis

remains on mechanical correctness and a certain amount of strategic repetition (White, 1994). Geisler (1995) found that teachers seem justified in not providing instruction for extended analytic writing, because few school subjects require such skills. Despite new directions in performance assessment, the classroom, in many cases, remains as that described by Applebee (1981) where the majority of students' writing exists in a testing context rather than an instructional one, and where few opportunities are provided to practice writing skills for which students will be formally assessed. This conflict between the notion of academic writing in the classroom and the skills we seek from writers in large-scale assessments is at the heart of needed study.

Several such studies have examined how large-scale assessments alter classroom practices or teacher perceptions. Baker, Freeman and Clayton in an important text, *Testing and Cognition*, (1991) examined the influence of large-scale essay testing in secondary school history classes. Much of their work focused on the connection between history content knowledge and the criteria by which to accurately assess that knowledge in student compositions. Torrance (1993) examined the influence of the national assessment in England and Wales piloted for first grade students. He discovered that rather than using the demands of performance assessment to alter instruction, teachers treated the assessment as a separate phenomenon and provided material and information only to prepare students for the exam. In surveys conducted by Koretz, Mitchell, Barron and Keith (1996) teachers in Maryland indicated that changes to classroom instruction were primarily made by sharing materials with students and in practicing for the assessment. Again, few significant changes in instruction were motivated by the use of performance assessments. Firestone and colleagues (1998) discovered similar results through research regarding mathematics performance assessments in Maine and Maryland. They concluded that teachers primarily gave students practice sessions and test preparation activities but did not make

substantial changes in methodology or classroom instruction. Fuchs, L. and colleagues (1999) studied the impact which performance assessments in mathematics classes had on teacher practice. Through extensive training with performance assessments, teachers did alter their strategies to provide more instruction aimed at problem-solving and analytical skills and less on basic concepts than teachers who did not receive intensive training. The importance of on-going and intensive training to bridge the connection between performance assessment and classroom practices is emphasized throughout these studies (Baker, Freeman & Clayton, 1991; Firestone, Mayrowetz & Fairman, 1998; Fuchs et al., 1998; Torrance, 1993).

Bridge and Hiebert (1985) conducted a preliminary study to determine the amount and type of writing that students did throughout the school day and the kind of writing instruction provided by teachers in elementary classrooms. They conducted teacher surveys to determine their self-perceptions of instructional practices by rating the frequency with which teachers used various writing activities. They concluded that student time spent on writing activities increased throughout the elementary grades. More importantly, however, they discovered that few of the writing practices which teachers employed taught students about the writing process or how to use writing strategies. In a follow-up study, Bridge and others (1997) studied the effects of the Kentucky Education Reform Act (KERA) of 1990 on teachers' writing instruction practices. Using classroom observations and teacher surveys, Bridge and her colleagues gauged the teachers' perceptions of the frequency in using specific writing tasks. In addition, the researchers sought to determine the teachers' perceptions of changes in practices as an outcome of KERA. The findings revealed that teachers were more inclined to promote higher-level writing strategies while decreasing their use of lower-level skills as a result of the state mandates. They found that teachers more frequently engaged students in crafting activities as well as the use of revision and

editing processes. The researchers concluded that many of these changes were directly attributable to the statewide assessment system.

Pennsylvania System of School Assessment

The Pennsylvania System of School Assessment (PSSA) provides comparative data to schools and districts for use in planning and informing the public. *The Writing Assessment Handbook* (1998-99) explains the goals and procedures used in the state writing assessment. The handbook states that, "The Pennsylvania Department of Education believes this direct or performance assessment of writing will enhance instructional efforts by encouraging teachers to design instruction around thoughtful, effective and meaningful writing tasks" (p.4). The Pennsylvania Department of Education urges districts to use the PSSA to inform teaching and to serve as a model for assessment development. Districts are encouraged to use the procedures and methods described in the assessment handbook to create local assessments in order to improve programs and instruction. The writing assessment procedures were established in 1989 by the Writing Assessment Advisory Committee (WAAC). In their efforts to design a writing test, the committee examined writing research and the various types of writing assessment used by several other states and the National Assessment of Educational Progress (NAEP).

The Pennsylvania writing assessment is given to all students in grades 6 and 9; recent legislation also requires the writing assessment in grade 11. Students respond to one of nine possible prompts or open-ended writing tasks that focus on one of three modes: narrative/imaginative, informational, or persuasive. Students get two forty-minute sessions to plan and compose their response. Students do not receive individual results on these writing exams. Instead, each district and school receives scaled scores of the achievement results and comparisons with the results of other similar schools.

The PSSA (1989) created the Pennsylvania Holistic Scoring Guide that serves as the scoring rubric for the state writing assessment. It includes scoring criteria along with five characteristics upon which the criteria are based: focus, content, organization, style, and conventions. State readers using this rubric holistically score student responses. Readers rate each written response on a scale of 1 to 6 based on the pre-established rubric and in comparison to anchor papers which serve as exemplars or models for responses at each score level. A paper's score is based on the total impression created by the piece in relation to the rubric and anchor papers. Each student's response is scored by at least two readers; the total score is a composite of the two readers' scores. If the two readers' scores are more than one point apart, a third reader acts as an arbitrator to determine the final score of the piece. Because one of the assessment's purposes is to provide districts with information on growth in writing for its students, the same scoring standards are used for all papers. For this reason, sixth and ninth grade papers are currently mixed together for the scoring. The scorers do not know what grade the writing samples are from, nor do they know the score given to any piece by another reader. The aim here is to apply the same standard to all writing.

To be truly meaningful, the writing assessments mandated by the Pennsylvania Department of Education must become significant to teachers in reconceptualizing their theoretical stance which in turn has immediate and on-going implications for their day-to-day writing instruction and classroom practices. While large-scale assessments such as the Pennsylvania writing assessment typically do not offer students opportunities to reflect upon results, they do permit districts to gauge the effectiveness of group scaled scores for their schools. The key factor is if and how individual classroom teachers use large-scale assessment results and the criteria by which they are determined to alter or modify their teaching strategies and classroom practices.

Methods

One method typically used to study the connection between large-scale assessment and teacher practices or attitudes is through survey design (Bridge & Hiebert, 1985; Fine, 1998). In this study, a cross-sectional comparison of Pennsylvania public secondary and elementary teachers was conducted. The survey was designed to examine teachers' perceptions regarding writing instruction, professional development activities, and their knowledge and use of the Pennsylvania writing assessment information and results. The field-tested survey contained sixty items. The first thirty items asked teachers to respond to a frequency scale regarding classroom practices directly related to the Pennsylvania Writing Assessment Holistic Scoring Guide. For example, teachers were asked how often they used the PSSA rubric to assess classroom writing assignments. Response choices were: "Almost every week," "Once or twice a month," "Once or twice a semester," "Once or twice a year," or "Never, or hardly ever." The second section of the survey asked teachers to respond to fifteen items on a level of agreement scale regarding teaching methodology. Teachers indicated their level of agreement with such statements as "The PSSA writing assessment has improved my ability to teach writing," by selecting one of these response choices: "Strongly agree," "Agree," "Uncertain," "Disagree," or "Strongly disagree." The final section of the survey requested demographic information regarding teacher experience and training along with school and district characteristics.

The survey sample was determined through a two-stage selection process. The selection of schools for participation in the survey was based on information from the Pennsylvania Department of Education. Schools for the survey were identified through an analysis of PSSA writing assessment scaled scores for two consecutive years, 1997 and 1998. Selection of schools

from that list was further defined by selecting only those schools that were at least thirty points above or below the similar schools band in both years. The state identifies a set of similar schools based on socioeconomic characteristics. In most cases this is the ten schools ranked immediately above and the ten schools ranked immediately below a school with respect to socioeconomic status. The similar schools band is found by taking the 20 similar schools in a set and ordering them according to the scaled score for writing. The scaled score by the sixth school from the bottom represents the lower limit of the similar schools band, while the score of the fifteenth school from the bottom represents the upper limit of the similar schools band. According to the state, this represents the average range within which schools tend to score when socioeconomic conditions are taken into account. For the purpose of this study, efforts were made to select an even number of schools with scores above and below the bands. Efforts were also made to select schools above and below the bands with approximately the same number of students participating in the assessment and with similar percentages of low-income students based on 1998 school information. The Pennsylvania Department of Education also provided this data. Further efforts were made to have an equal number of elementary and secondary teachers respond to the survey. Superintendents of districts identified through this process were then contacted for permission to contact teachers. Surveys were mailed to 266 secondary English or elementary teachers responsible for language arts instruction. A total of 168 teachers (63%) from twenty schools in fifteen Pennsylvania public school districts responded to the survey.

Limitations of the Study

As with most research, the current study has limitations. First, the survey items are connected directly to the characteristics of effective writing and holistic scoring criteria and descriptors detailed in the state's scoring guide. In the same manner, the survey items focus on

the modes of writing and the sample materials described and provided as part of the state writing assessment. By concentrating specifically on these skills and details, many factors of writing instruction, such as revision processes and the use of journals and portfolios, could not be considered. The study is further limited to teachers in public schools who are responsible for writing instruction. This eliminates any input regarding cross-curricular or writing across the curriculum efforts.

The selection of schools based on PSSA writing scores also limits the input of teachers from all districts across the state. Of even more significance, the small number of respondents from certain response categories, for example few teachers responding to the survey were from very large or very small school districts, limits the interpretation of some results. In addition, of course, frequency responses alone do not capture the effectiveness of those classroom practices described by respondents.

The survey is limited also by relying on subjective teacher self-reporting of perceptions and practices rather than any concrete system of accounting for teacher behaviors or attitudes. Likewise the amount and content of classroom writing and instruction taking place prior to the state assessment implementation is difficult, if not impossible, to discern through this initial study.

These limitations in measurement, however, are counter-balanced by efforts to attain broad and balanced representation as described above and by a reliance on teacher, school, and district anonymity in the survey process and result reports and findings.

Results

Descriptive data and survey results are presented in three main sections: (a) Demographics, (b) Classroom Practices and (c) Teaching Methodology. Classroom Practices data, which summarizes teacher responses to thirty items on a 5-point frequency scale, is

sub-divided into three areas as well. These three areas are: Writing Experiences and Opportunities, Scoring Criteria Based on the Pennsylvania Holistic Scoring Guide, and Characteristics of Effective Writing. Response choices on the survey to these sections were coded so that the higher the score the more frequent the practice, i.e., a score of 1 equals "Never, or hardly ever"; a score of 5 equals "Almost every week." Finally Teaching Methodology, which summarizes the level of agreement by respondents to fifteen items, is presented in one section. Response choices were coded so that the greater the level of agreement with a statement, the higher the score. For example, a score of 1 equals the response "Strongly disagree"; a score of 5 equals "Strongly agree."

Demographics

Teachers responded to fifteen items to describe their experiences, responsibilities and training. Of the 168 respondents, teachers indicated that 56.9% teach elementary school or Grades 1-6; 43.1% indicated that they are secondary teachers for Grades 7-12. The majority of teachers (71.9%) have received training in holistic scoring; 27.5% of the respondents had not received training in holistic scoring.

An additional significant factor in the selection of schools participating in the survey was the state writing scores during two consecutive years, 1997 and 1998. Only those schools who scored at least 30 points "above" or "below" the similar schools band in both of those consecutive years were eligible to participate in the survey. The Pennsylvania Department of Education provided this data. Additional effort was made to balance school participation so that a fairly equal number of teachers from schools with scores above and below the similar schools band would be represented. Teacher responses revealed that 54.8% of them had school writing scores above the similar schools band during both the 1997 and 1998 years; 45.2% of the respondents

came from schools that scored below the similar school band on the state writing assessment in both the 1997 and 1998 years.

Classroom Practices

Teachers were asked to respond to thirty items regarding classroom practices. The items were designed to solicit input on the writing experiences which teachers provide students and how those experiences relate to the skills and demands of the state writing assessment. In addition, several items required teachers to reflect on if and how they used the Pennsylvania Writing Assessment Holistic Scoring Guide, as well as other training materials and writing samples provided by the state, as part of their classroom instruction. Teachers responded to a Likert scale on these thirty items to measure their frequency in using these practices. Frequency response choices to these statements were: "Never, or hardly ever," "Once or twice a year," "Once or twice a semester," "Once or twice a month," or "Almost every week."

Of the three modes of discourse used in the Pennsylvania Writing Assessment, (Narrative/Imaginative, Informative, and Persuasive) the mode least likely to be practiced by students on a weekly basis is in response to persuasive topics at 14.3%. That low frequency is consistent with research findings that show persuasive prompts to be the most challenging, even for skilled writers who have difficulty focusing or organizing their responses (Masters, 1992).

The overall results to these items seem to indicate that less than half of the teachers who responded to the survey consistently use -- that is "Once or twice a month" or more -- the scoring criteria and rubrics from the Pennsylvania Holistic Scoring Guide. While teachers use rubrics and provide instruction about holistic scoring, they are not very likely to use the state scoring guide or samples from the PSSA materials to illustrate the rubric or explain holistic scoring in their own classroom instruction.

The Pennsylvania Writing Assessment Holistic Scoring Guide establishes five characteristics of effective writing: focus, content, organization, style, and conventions. The guide also offers three to four descriptors for each of these characteristics as the basis for the holistic scoring. Focus, for example, is explained with four descriptors. They include: "demonstrates an awareness of audience and task," "establishes and maintains a clear purpose," "sustains a single point of view," and "exhibits clarity of ideas." PSSA training materials further illustrate these characteristics through samples of student writing. Teachers who responded to the survey were asked sixteen items which focused specifically on these five characteristics as they relate to their classroom instruction.

For each of the five characteristics of effective writing, it is interesting to note that the characteristic itself received importance in classroom instruction, while the descriptors provided in the state's holistic scoring guide play much less importance - and the samples from PSSA, even less importance - in classroom instruction. This pattern holds true for each characteristic. For example, while 49.4% of teachers say they explain the importance of focus (Focus) "Almost every week," only 8.3% indicate that they use the descriptors of focus (Focus Descriptors) from the Pennsylvania scoring guide that often. In addition, an even smaller percentage of 3.6% say they use the samples of focus (Focus Samples) from the PSSA materials that frequently. The use of samples to illustrate each characteristic received the largest response rates in the "Never, or hardly ever" level: Organization Samples (52.4%), Style Samples (52.4%), Convention Samples (50.6%), Focus Samples (50.3%), and Content Samples (50.0%). The responses reveal that teachers overall place emphasis on the five characteristics in classroom instruction of student writing, but do not use the Pennsylvania Holistic Scoring Guide descriptors or the PSSA samples to explain or illustrate them.

Teaching Methodology

Teachers were asked to respond to fifteen survey items about their beliefs regarding the state writing assessment in relation to their teaching methods. The items required teachers to respond to a Likert scale for level of agreement with the statement. Response choices were "Strongly disagree," "Disagree," "Uncertain," "Agree," and "Strongly agree."

The overall responses indicate that teachers have a general level of agreement with the beliefs and influences of the state writing assessment in relation to their teaching methodology. They agree that the five characteristics of effective writing, supported through the state holistic scoring guide, are consistent with their beliefs, and that they are thoroughly familiar with the criteria in the same guide. The strongest level of agreement, however, is the importance placed on the results by the district. Teachers feel that their approach to teaching writing has not been significantly altered by the state assessment and that it has not improved their ability to teach writing. Likewise they do not see the Pennsylvania scoring guide as a valuable teaching resource.

Data Analysis

While the initial study examined the effect of sixteen factors on classroom practices and teaching methodology, this paper focuses on three key factors: Teaching Responsibility (i.e. elementary vs. secondary teachers) Holistic Training, and PSSA Scores above or below the similar schools band. The following tables summarize the effects of these three factors. The classroom practice variables are sub-divided into three tables: Writing Experiences and Opportunities, Scoring Criteria Based on the Pennsylvania Holistic Scoring Guide, and the Characteristics of Effective Writing. Teaching Methodology items are presented in one table.

For each separate series of related t-tests conducted in these tables, the alpha level was adjusted to detect more specific significance levels in accordance with the Bonferoni suggestion

(Rosenthal & Rosnow, 1984). In Table 1 the alpha level was adjusted to detect significance only for $p<.005$ (.05/9); Table 2 alpha levels were adjusted to detect $p<.01$ (.05/5); the alpha levels for Tables 3 and 4 were adjusted to detect a significance only of $p<.003$ (.05/16 and .05/15).

Classroom Practices – Writing Experiences and Opportunities

Nine items asked teachers to indicate how frequently they provided their students with classroom writing experiences and opportunities directly related the state assessment. As Table 1 demonstrates, there is no significant difference between the mean frequencies of the experiences provided between elementary and secondary teachers. Likewise, there is no difference between teachers from schools who performed above or below the similar schools band. Two practices reflect significant differences between teachers who have been trained with those who have not been trained. Teachers who have received holistic training are more likely to use holistic scoring (Holistic Scoring) for classroom writing assignments and more frequently provide students timed writing assignments (Timed Writing) as part of classroom instruction.

While there are not significant differences between elementary and secondary teachers in the use of these classroom practices, mean scores reveal that each group most frequently provides students with opportunities to write based on impromptu topics and in response to narrative/imaginative prompts. The mean scores for those practices indicate that teachers provide those opportunities approximately once or twice a month. All teachers are least likely to give students sample prompts or to conduct practice sessions for the state writing assessment; means scores reflect that teachers employ those practices approximately once or twice a year.

Means scores also show that teachers in schools that perform well on the state assessment as well as those teachers in schools that perform below the similar schools band are more likely to have students write in response to impromptu topics at least once or twice a month. All teachers

Table 1.

Summary of Teaching Responsibility Effects/Holistic Training/PSSA Scores on Classroom Practices - Writing Experiences and Opportunities

Variable		Teaching Responsibility		t	Holistic Training		t	PSSA Scores-Similar Schools Band		t
		Elementary (n = 93-95)	Secondary (n = 71-72)		Yes (n=117-129)	No (n=45-46)		Above (n=91-92)	Below (n=74-76)	
Sample Prompts	M	2.65	2.52	0.62	2.77	2.20	2.56	2.55	2.65	-0.46
	SD	(1.46)	(1.11)		(1.30)	(1.29)		(1.39)	(1.22)	
Impromptu Writing	M	4.12	3.88	1.50	4.04	3.90	0.84	3.91	4.13	-1.38
	SD	(4.12)	(3.88)		(1.02)	(1.04)		(1.09)	(0.88)	
Narr./Imaginative	M	4.03	3.77	1.59	3.90	3.98	-0.40	3.91	3.92	-0.04
	SD	(1.00)	(1.06)		(1.00)	(1.02)		(1.09)	(0.97)	
Informational	M	3.95	3.65	1.78	3.86	3.72	0.77	3.77	3.86	-0.50
	SD	(1.04)	(1.09)		(1.06)	(1.05)		(1.13)	(1.00)	
Persuasive	M	3.20	3.32	-0.64	3.30	3.20	0.51	3.22	3.29	-0.39
	SD	(1.29)	(1.03)		(1.16)	(1.22)		(1.20)	(1.16)	
Open-ended	M	3.58	3.54	0.19	3.70	3.26	1.97	3.46	3.69	-1.14
	SD	(1.35)	(1.27)		(1.22)	(1.45)		(1.48)	(1.06)	
Timed Writing	M	2.55	2.81	-1.10	2.90	2.09	3.23*	2.79	2.47	1.38
	SD	(1.51)	(1.46)		(1.47)	(1.38)		(1.59)	(1.34)	
Holistic Scoring	M	2.71	2.97	-1.11	3.12	2.13	3.82*	2.79	2.88	-0.36
	SD	(1.58)	(1.45)		(1.48)	(1.44)		(1.59)	(1.46)	
Practice Sessions	M	2.75	2.39	1.57	2.77	2.24	2.12	2.74	2.43	1.34
	SD	(1.59)	(1.25)		(1.47)	(1.34)		(1.60)	(1.24)	

* $p < .005$

are less likely – once or twice a year - to give students sample prompts from the state assessment or to conduct practice sessions.

Examining the means scores of teachers who have received training with those who have not been trained reveals a similar pattern. The classroom practices most likely employed, as shown through mean scores, are opportunities to write in response to impromptu topics and in response to narrative/imaginative prompts. Means scores indicate that all teachers use those practices in their classrooms nearly once or twice a month. The use of sample prompts and practice sessions are employed only once or twice a semester or less.

It should be noted that the survey items regarding the use of open-ended prompts, timed writing assessments and holistic scoring experiences asked teachers for frequency rates in the use of those practices prior to implementation of the state writing assessment in their respective districts. In other words, teachers who had been trained in holistic scoring indicated that, prior to the state assessment in their districts, they had provided students with timed writing performance assessments ($\underline{M}$=2.90) and holistic scoring experiences ($\underline{M}$=3.12) at least once or twice a semester. Untrained teachers responded that they provide those experiences – timed writing assignments ($\underline{M}$=2.09) and holistic scoring ($\underline{M}$=2.13) – closer to once or twice a year.

Classroom Practices – Scoring Criteria Based on the PSSA Holistic Scoring Guide

Five items on the survey asked teachers the frequency of classroom practices directly related to use of the scoring criteria from the state holistic scoring guide. Teachers were asked how often they explain the state scoring criteria and how frequently they use the six-point scale and samples from PSSA materials in their classrooms. In addition they were asked to indicate how often they instruct students about holistic scoring and use rubrics for classroom assessments.

Reflected in Table 2, teaching responsibility has no significant effect on classroom practices related to use of the scoring criteria in the Pennsylvania Holistic Scoring Guide. The means scores show that all teachers are least likely to use samples from the PSSA materials to illustrate the six-point scale; elementary teachers (M=2.13) and secondary teachers (M=2.10). Both elementary and secondary teachers use rubrics for classroom assessment more frequently-greater than once or twice a semester; elementary (M=3.39) and secondary teachers (M=3.51).

There is significant difference in the use of scoring criteria in classrooms between teachers in schools that scored above or below the similar schools band. As the means scores show, teachers in schools that score below the similar schools band (M=3.14) share and explain the scoring criteria of the state assessment with students as part of classroom instruction once or twice a semester. Teachers in higher achieving schools (M=2.49) share the state assessment scoring criteria approximately once or twice a year. The higher frequency rates from teachers in schools that score below the similar schools band for those practices may indicate that teachers are attempting to more frequently address the issue in classrooms in an effort to improve writing assessment scores. The frequencies of these classroom practices, for the most part, indicate that there are no differences between teachers in high or low achieving schools. The mean scores, however, may reveal an effort by teachers in schools who have not performed well to increase instruction relevant to the state scoring criteria and holistic scoring guide.

Holistic training effects are evident for each of the five classroom practices related to the scoring criteria of the holistic guide. Teachers who have been trained in holistic scoring are significantly more likely to utilize each of these classroom practices than teachers who have not been trained. Trained teachers utilize these classroom practices slightly more than once of twice a semester, while untrained teachers employ those practices no more than once or twice a year.

Table 2. Summary of Teaching Responsibility / Holistic Training / PSSA Score Effects on Classroom Practices – Scoring Criteria Based on the Pennsylvania Writing Assessment Holistic Scoring Guide

		Teaching Responsibility			Holistic Training			PSSA Scores –Similar Schools Band		
Variable		Elementary (n = 95)	Secondary (n = 71-72)	t	Yes (n=119-120)	No (n=46)	t	Above (n=92)	Below (n=73-76)	t
Scoring Criteria	M	2.81	2.76	0.21	3.04	2.13	3.77*	2.49	3.14	-3.00*
	SD	(1.58)	(1.26)		(1.43)	(1.26)		(1.45)	(1.36)	
Six-point Scale	M	2.83	2.72	0.46	3.13	1.93	4.70*	2.53	3.12	-2.45
	SD	(1.46)	(1.68)		(1.52)	(1.34)		(1.61)	(1.45)	
Six-point Samples	M	2.13	2.10	0.13	2.35	1.54	3.66*	2.16	2.05	0.53
	SD	(1.39)	(1.24)		(1.34)	(1.07)		(1.45)	(1.14)	
Holistic Instruction	M	3.01	3.10	-0.42	3.39	2.24	5.31*	3.02	3.11	-0.40
	SD	(1.44)	(1.21)		(1.23)	(1.29)		(1.44)	(1.23)	
Rubrics	M	3.39	3.51	-0.53	3.76	2.65	4.82*	3.29	3.64	-1.63
	SD	(1.42)	(1.44)		(1.25)	(1.52)		(1.52)	(1.28	

* $p<.01$

The general mean scores of these classroom practices show that all teachers are more likely to share the state scoring criteria with their students and use rubrics to assess classroom writing assignments. They are less likely to use the state six-point scale or to illustrate the scale with samples from the PSSA materials.

Classroom Practices – Characteristics of Effective Writing

The Pennsylvania Holistic Scoring Guide details five characteristics of effective writing: focus, content, organization, style and conventions. Sixteen survey items solicited the frequency of classroom practices directly related to these characteristics and the supporting materials and information provided by the state to illustrate them. Table 3 summarizes the effect of Teaching Responsibility, Holistic Training, and PSSA Scores on those sixteen classroom practices.

As the table demonstrates, there are no significant differences in the frequency of these practices between elementary and secondary teachers. There are also no significant differences between the practices of teachers in schools that score above the similar schools band with those of teachers in schools that score below the similar schools band.

Teachers responses do reveal significant differences between teachers who have received holistic training with those who have not been trained for six of these classroom practices. Teachers who have been trained in holistic scoring share the five characteristics with their students at least once or twice a semester. Those who have not received training share the characteristics with their students only once or twice a year. For each of the specific characteristics, trained teachers are more likely to use the descriptors from state holistic scoring guide to illustrate the characteristics than those teachers who have not had training. Mean scores reveal that untrained teachers use the guide descriptors for each characteristic less than once or twice a year.

Table 3. Teaching Responsibility/Holistic Training/PSSA Score Effects on Classroom Practices - Characteristics of Effective Writing

Variable		Teaching Responsibility			Holistic Training			PSSA Scores- Similar Schools Band		
		Elementary (n=93-95)	Secondary (n=70-72)	t	Yes (n=117-120)	No (n=46)	t	Above (n=92)	Below (n=73-76)	t
Share Five Characteristics	M	2.76	2.90	-0.65	3.11	2.11	4.21*	2.74	2.95	-0.92
	SD	(1.46)	(1.43)		(1.41)	(1.27)		(1.55)	(1.31)	
Focus	M	4.20	4.21	-0.06	4.30	3.96	2.05	4.23	4.18	0.29
	SD	(1.07)	(0.84)		(0.87)	(1.19)		(1.05)	(0.88)	
Focus Descr	M	2.60	2.51	0.40	2.85	1.93	3.40*	2.42	2.76	-1.59
	SD	(1.45)	(1.28)		(1.36)	(1.22)		(1.51)	(1.20)	
Focus Samp	M	2.11	2.04	0.32	2.28	1.65	2.86	2.17	1.99	0.93
	SD	(1.36)	(1.20)		(1.32)	(1.12)		(1.41)	(1.12)	
Content	M	4.21	4.26	-0.32	4.33	4.04	1.87	4.30	4.15	1.05
	SD	(0.99)	(0.83)		(0.77)	(1.11)		(0.99)	(0.82)	
Content Descr	M	2.43	2.33	0.47	2.70	1.67	4.65*	2.32	2.50	-0.88
	SD	(1.40)	(1.28)		(1.34)	(1.08)		(1.44)	(1.23)	
Content Samp	M	2.19	1.94	1.23	2.23	1.80	1.89	2.17	2.00	0.86
	SD	(1.39)	(1.16)		(1.29)	(1.29)		(1.40)	(1.17)	
Organization	M	4.41	4.46	-0.36	4.48	4.26	1.52	4.41	4.45	-0.26
	SD	(0.89)	(0.79)		(0.82)	(0.91)		(0.92)	(0.76)	
Org. Descr.	M	2.44	2.35	0.40	2.68	1.76	3.90*	2.28	2.56	-1.26
	SD	(1.49)	(1.31)		(1.40)	(1.25)		(1.46)	(1.35)	
Org. Samp.	M	2.14	1.99	0.74	2.28	1.61	3.03	2.11	2.05	0.27
	SD	(1.36)	(1.25)		(1.35)	(1.08)		(1.39)	(1.23)	
Style	M	3.75	3.92	-0.89	3.99	3.43	2.73	3.85	3.79	0.31
	SD	(1.25)	(1.17)		(1.10)	(1.36)		(1.30)	(1.10)	
Style Descr.	M	2.33	2.21	0.55	2.54	1.67	3.80*	2.27	2.30	-0.15
	SD	(1.43)	(1.28)		(1.35)	(1.21)		(1.48)	(1.22)	
Style Samp.	M	2.18	1.85	1.66	2.20	1.70	2.27	2.12	1.96	0.79
	SD	(1.39)	(1.12)		(1.30)	(1.23)		(1.39)	(1.17)	
Conventions	M	3.78	4.11	-1.59	4.10	3.50	2.67	3.91	3.93	-0.10
	SD	(1.43)	(1.19)		(1.18)	(1.56)		(1.41)	(1.25)	
Conv. Descr.	M	2.55	2.40	0.63	2.78	1.80	4.01*	2.39	2.62	-1.00
	SD	(1.55)	(1.37)		(1.46)	(1.26)		(1.50)	(1.43)	
Conv. Samp.	M	2.16	1.97	0.91	2.27	1.67	2.64	2.12	2.05	0.32
	SD	(1.40)	(1.21)		(1.36)	(1.16)		(1.41)	(1.22)	

* $p < .003$

These results indicate that all teachers make instruction of the characteristics of effective writing part of classroom instruction. They are not likely, however, to depend heavily on the Pennsylvania Holistic Scoring Guide or the state materials as part of that instruction. Mean scores demonstrate that all teachers devote classroom time to discussing the characteristics of effective writing. Most frequently – more than once or twice a month – all teachers explain the importance of focus, content, and organization in student writing. Slightly less frequently – at least once or twice a semester – they explain the importance of style and conventions. Also apparent from the mean values for these frequencies is that all teachers are more likely to explain the importance of each characteristic and less likely to use the descriptors provided in the Pennsylvania Holistic Scoring Guide. Teachers are even less inclined to use PSSA samples as part of classroom instruction.

Teaching Methodology

The final fifteen survey items focused on teaching methodology. Teachers were asked to indicate their level of agreement with principles and beliefs associated with the state assessment. Table 4 summarizes the effect of Teaching Responsibility, Holistic Training and PSSA Scores on these variables. The higher the mean value, the greater the level of agreement with the statement.

Two of the fifteen items reflect significant differences in teaching methodology between elementary and secondary teachers. Secondary teachers ($\underline{M}$=4.11) feel more strongly that they are thoroughly aware of the scoring criteria from the Pennsylvania Holistic Scoring Guide than elementary teachers ($\underline{M}$=3.49). Secondary teachers ($\underline{M}$=4.08) also more strongly agree that they are aware of the process by which the state writing assessment is conducted in their districts than elementary teachers ($\underline{M}$=3.40).

Table 4. Summary of Teaching Responsibility/Holistic Training/PSSA Score Effects on Teaching Methodology

Variable		Teaching Responsibility		t	Holistic Training		t	PSSA Scores-Similar Schools Band		t
		Elementary (n=94-95)	Secondary (n=71-72)		Yes (n=119-120)	No (n=46)		Above (n=91-92)	Below (n=76)	
Teaching Approach	M	3.57	3.08	2.57	3.51	2.98	2.52	3.29	3.43	-0.74
	SD	(1.22)	(1.20)		(1.20)	(1.24)		(1.26)	(1.18)	
Review Results	M	3.45	3.54	-0.50	3.69	2.98	3.74*	3.41	3.59	-1.02
	SD	(1.23)	(1.02)		(1.03)	(1.27)		(1.19)	(1.07)	
Material Influential	M	3.27	3.39	-0.64	3.55	2.78	4.05*	3.21	3.47	-1.52
	SD	(1.22)	(1.03)		(1.05)	(1.19)		(1.17)	(1.09)	
Valuable Resource	M	2.85	3.21	-1.86	3.28	2.30	4.82*	2.80	3.26	-2.44
	SD	(1.30)	(1.13)		(1.18)	(1.11)		(1.27)	(1.15)	
Altered Teaching	M	3.47	3.39	0.46	3.66	2.88	4.04*	3.26	3.66	-2.21
	SD	(1.24)	(1.09)		(1.09)	(1.22)		(1.26)	(1.03)	
Holistic Beliefs	M	3.34	3.71	-2.22	3.68	3.04	3.50*	3.35	3.70	-2.09
	SD	(1.18)	(0.88)		(1.00)	(1.15)		(1.20)	(0.88)	
Characteristic Beliefs	M	3.78	4.15	-2.74	4.06	3.67	2.52	3.88	4.03	-1.06
	SD	(1.02)	(0.64)		(0.89)	(0.87)		(1.03)	(0.69)	
Improved Teaching	M	3.19	3.04	0.83	3.29	2.70	3.09*	2.90	3.41	-2.94
	SD	(1.15)	(1.12)		(1.14)	(1.03)		(1.16)	(1.05)	
Improved Ability	M	3.17	3.23	-0.33	3.35	2.78	3.07	2.96	3.49	-3.21*
	SD	(1.22)	(0.91)		(1.08)	(1.05)		(1.14)	(0.96)	
Improved Skills	M	3.20	2.94	1.56	3.28	2.60	3.79*	2.96	3.23	-1.88
	SD	(1.17)	(0.88)		(1.04)	(0.96)		(1.07)	(1.02)	
District Results	M	4.01	4.00	0.07	4.12	3.76	2.17	4.26	3.71	3.88*
	SD	(1.00)	(0.89)		(0.88)	(1.10)		(0.80)	(1.04)	
Scoring Awareness	M	3.49	4.11	-3.54 *	4.08	2.98	6.02*	3.76	3.78	-0.09
	SD	(1.25)	(0.90)		(0.98)	(1.22)		(1.22)	(1.08)	
Process Awareness	M	3.40	4.08	-4.18 *	3.93	3.11	4.52*	3.73	3.67	0.36
	SD	(1.21)	(0.80)		(0.93)	(1.32)		(1.16)	(1.04)	
Student Results	M	3.41	3.42	-0.04	3.59	2.98	3.45*	3.54	3.28	1.65
	SD	(1.07)	(1.03)		(1.05)	(0.95)		(1.04)	(1.05)	
District Training	M	3.21	3.17	0.26	3.48	2.50	5.50*	3.33	3.05	1.61
	SD	(1.11)	(1.09)		(1.05)	(0.94)		(1.06)	(1.14)	

* $p < .003$

Table 4 also reflects significant differences between teachers in high achieving schools with those in low achieving schools for two teaching methodology items. Teachers in schools that scored below the similar schools band (M=3.49) agree more strongly than those in schools above the band (M=2.96) that the Pennsylvania writing assessment has improved the ability of students to write in response to open-ended prompts. Teachers in higher achieving schools may believe that their daily instruction and consistent practices, more than the state assessment itself, accounts for improved student responses to open-ended assignments. Teachers in lower performing schools may see the state assessment as the key factor toward improving student responses to open-ended prompts. Their agreement with the impact of the state assessment may have motivated more open-ended writing tasks and a greater awareness of the influences the state assessment has on student skills.

Teachers in higher achieving schools (M=4.26) agree more strongly than those in lower achieving school (M=3.71) that their districts consider it important to share the results of the state assessment with teachers. This could demonstrate that districts that share results offer teachers more reflection on those results and demonstrate a greater commitment of those districts toward writing achievement. It could also reflect an earlier concern by those districts to the state assessment of writing which motivated teachers to focus on those skills and strategies associated with holistic scoring and large-scale, state assessment efforts.

The effects of training on teacher methodology are clearly evident. Eleven of the fifteen items demonstrate significant differences in the beliefs and values of teachers trained in holistic scoring with those of teachers who have not been trained. Panel one indicates that holistic training significantly effects how strongly teachers review and reflect upon assessment results, how much influence state materials are in the classroom, how valuable the holistic scoring guide is

to teaching writing and how the state assessment has altered teaching methods. Teachers who have been trained more strongly agree with the influences and impact that the state assessment has had on these values and beliefs. Untrained teachers disagree to a greater extent with the influences and impact the state assessment has had on their teaching and beliefs.

Panel two shows that trained teachers (M=3.68) agree more strongly that holistic scoring is consistent with their beliefs about writing instruction than teachers who have not been trained (M=3.04). Trained teachers also feel more strongly that the state assessment has improved their teaching (M=3.29) than teachers who have not had holistic scoring training (M=2.70). Likewise, trained teachers (M=3.28) have a greater level of agreement than untrained teachers (M=2.60) that the state assessment has improved the overall writing skills of their students.

Panel three indicates that those teachers trained in holistic scoring significantly differ from teachers who have not been trained in their level of agreement with four items. They agree more strongly that they are aware of scoring criteria from the Pennsylvania Holistic Scoring Guide and the process used to conduct the assessment in their district. Trained teachers also feel more strongly that their districts makes students aware of the assessment results and have offered training on preparing students for the writing assessment. Untrained teachers disagree with or are uncertain about the scoring criteria and processes as well as the importance placed by their districts on sharing results or on training aimed as preparing students for the state assessment.

Discussion

Teacher responses to the survey indicate that they are providing students with frequent writing experiences and opportunities. Teachers are most likely to give students opportunities to write in response to impromptu topics and in the narrative/imaginative and informational modes. Also revealed in the study, however, is the reluctance of teachers overall to fully utilize timed-

writing and holistic scoring experiences. Teachers do not invest themselves and their classroom instruction in the scoring criteria and samples generated as part of the state assessment. The results also indicate that teachers value the five characteristics of effective writing as presented in the Pennsylvania Holistic Scoring Guide and make instruction about those characteristics important parts of classroom instruction. Again, however, they are reluctant to base that classroom instruction on materials created by the state.

Several significant findings about factors that influence classroom practices are also revealed in the study. First, teaching responsibility, that is the difference between elementary and secondary teachers, does not significantly influence the use of any of classroom practices associated with the Pennsylvania writing assessment. While few studies have focused on teaching responsibility effects, these results are supported by the work of Koretz and colleagues (1996) who examined the effects of the Kentucky Instructional Results Information System (KIRIS). They found only limited differences between teachers in grades six and eight in response to assessment reform efforts.

Also showing little influence on these practices are PSSA scores above or below the similar schools band. Only one classroom practice reflects significant differences between teachers in high achieving and low achieving schools. Teachers in schools that scored below the similar schools band were more likely to share the scoring criteria from the state holistic scoring guide with their students. The fact that such limited, significant difference exists regarding these practices and actual PSSA writing assessment scores speaks to the complex nature of writing instruction and writing performance.

The key factor influencing classroom practices is teacher training. Training significantly affects the value which teachers place on writing assessments and how frequently they utilize

teaching strategies connected directly to the skills and demands of the state writing assessment. The impact of holistic training is also evident in teaching methodology. Teachers who have received training significantly differ in their beliefs with those teachers who have not had training. For the majority of teaching methodology items, those teachers who have been trained in holistic scoring more strongly agree that the state writing assessment has influenced them and improved their ability to teach writing. In addition, they concur with the beliefs of holistic scoring and the five characteristics of effective writing. They also agree more strongly that the state test has improved the writing ability and skills of their students as well as their own ability to teach writing. They agree that their students have stronger awareness of the district writing results and that they have greater awareness of the scoring criteria and processes used in their districts to conduct the state assessment. These results emphasizing the importance of training on teaching practices and beliefs duplicate that of other research (Bridge et al., 1997; Fuchs et al., 1999; Torrance, 1993).

Based on survey results and analysis, several key findings emerge. First, the state assessment seems to have motivated teachers to present numerous writing experiences and opportunities to students in both the elementary and secondary schools. Secondly, teachers recognize the value of holistic scoring and the characteristics of effective writing as presented in the Pennsylvania Holistic Scoring Guide. In addition, teachers include those characteristics of effective writing as part of classroom instruction. Most clearly, those factors related to teacher training in holistic scoring are significant determinants in the frequency of classroom practices directly related to the skills and demands of the state large-scale assessment and with the values and procedures for holistic scoring. It is evident that even though many teachers agree with the beliefs and values of holistic scoring and the characteristics of effective writing, they are reluctant

to fully use the state rubric, descriptors and writing samples as the basis for classroom instruction of those skills and characteristics. This data may suggest a weakness in the supporting materials provided by the state. It may indicate instead that teachers are designing and providing students with their own rubrics and samples to detail and illustrate the characteristics of focus, content, organization, style and conventions. It may also be, however, that teachers are teaching those characteristics while clinging to traditional methods and not totally adopting a holistic or progressive approach to writing instruction. The possibility that teachers are making students aware of the rubrics and characteristics of effective writing without concerted instructional activities to increase student appreciation or understanding is consistent with other studies (Firestone, Mayrowetz & Fairman, 1998; Koretz, Mitchell, Barron & Keith, 1996; Torrance, 1993).

Of greater importance is that reliance on and use of state materials and processes related to the state assessment do not in themselves necessarily result in improved state writing assessment scores. In other words, a dependency on classroom usage of the state rubric, support materials and holistic scoring procedures does not automatically translate into improved assessment scores on the state writing exam. It must be seen then that state policies and procedures which focus on holistic training and scoring are not a panacea toward the overall improvement of writing skills or writing performance. The complex nature of writing skill development and of writing instruction cannot be overlooked. Likewise, a short term, system-based solution mandated by state policy cannot be over-emphasized at the expense of other strategies and factors.

Yet there is certainly much to be said for conducting large-scale writing assessments. There is great potential to inspire change if we are willing to reason to the root of the results, to

boil down the numbers, and to get to the heart of the matter. Large-scale state writing assessments can provide valuable learning experiences for both students and teachers. Such assessments, as have been demonstrated, are valid indicators of student skills (Soltis & Walberg, 1989; Walberg & Ethington, 1991) and teacher and program effectiveness (Farr & Jongsma, 1993). Holistic assessment is also consistent with current pedagogy on language arts and composition (White, 1994). Ideally, large-scale assessments, like those used within the context of the classroom, should be designed to allow all stakeholders the chance to reflect upon the results and to make needed changes in instruction and curriculum to fortify strengths and to address weaknesses. Large-scale writing assessments, such as those mandated in Pennsylvania, can help bring about change only if strong efforts are made to use the results to inform and alter classroom writing instruction and to unify teachers toward those goals. Obviously intensive and on-going performance assessment training is essential if classroom practices are to be changed and more closely aligned with the goals and demands of large-scale writing assessments.

The potential influence of state mandated, large-scale writing assessment policies remains to be seen. How the state continues to shape and support training along with classroom related strategies and innovation will ultimately determine policy influence and effectiveness. To truly bring about positive changes in classroom instruction and student writing ability, those connections to state assessment policies must be more fully explored, addressed and solidified.

Conclusion

This study provides a baseline for future explorations regarding the many influences which state performance assessment policies have on classroom instructional practices and teaching methodology. Based on these current findings, future studies need to focus closely on training factors to better discern how training shapes classroom writing instruction beyond frequency

patterns. In addition to a focus on training, future studies will also need to include an examination of other instructional strategies that might impact writing achievement and results. Those factors should include such strategies as the writing process, portfolios, journals, graphic organizers, revision techniques, peer group editing, writing for publication, formal essay construction and course content knowledge.

This study was limited specifically to those characteristics, modes, and dimensions described and detailed in the Pennsylvania Writing Assessment Holistic Scoring Guide and state related materials. The survey and research were limited to items that directly connect to those characteristics and scoring procedures related to the performance assessment policies and procedures. Future studies should seek to go beyond those details for a fuller description of the practices and strategies writing teachers implement in their classrooms. Continued study should be undertaken into how such assessment policies motivate changes in attitudes and perspectives about writing, about how writing is taught, and about the role writing instruction plays in the overall philosophy and goals of teachers, schools and districts. In addition, how the state assessment results are presented, perceived and utilized by classroom teachers in their instructional practices must also be further explored. These are broad and sweeping concerns, yet they are significant and relevant to all schools and teachers who understand the importance of student writing skills as fundamental to the learning process.

References

AERA position statement concerning high-stakes testing in prek-12 education. (July 2000). [On-line]. Available: aera.net.about.policy.stakes.htm.

Applebee, A.N. (1981). A study of writing in secondary schools. Urbana, IL: National Council of Teachers of English.

Atwell, N. (1987). In the middle: Writing, reading and learning with adolescents. Portsmouth, NH: Boynton/Cook.

Baker, E. L., Freeman, M. & Clayton, S. (1991). In M.C. Wittrock & E.L. Baker (Eds.) Testing and cognition (pp. 131-165). Englewood Cliffs, NJ: Prentice Hall.

Barnes, D. (1975). From communication to curriculum. New York: Penguin.

Berthoff, A.E. (1982). Forming, thinking, writing: The composing imagination. Portmouth, NH: Boynton/Cook.

Bridge, C.A., & Hiebert. E.H. (1985). A comparison of classroom writing practices, teachers' perceptions of their writing instruction, and textbook recommendations of writing practices. Elementary School Journal, 86, 155-172.

Bridge, C.A., & others (1997). Classroom writing practices revisited: The effect of statewide reform on writing instruction. Elementary School Journal, 98, 151-170.

Britton, J. (1970). Language and learning. New York: Penguin.

Farr, R. & Jongsma, E. (1993). The convergent / discriminant validity of integrated reading/writing assessment. Journal of Research and Development in Education 26 (2), 83-91.

Fine, T. (1998). The impact of the Connecticut mastery test on teaching practices in a large suburban Connecticut town (Doctoral dissertation, University of Sarasota, 1998). ProQuest Digital Dissertations, AAT, 9824814.

Finn, C.E. & Petrilli, M. J. (2000) The state of state standards: 2000. Washington, D C: The Thomas B. Fordham Foundation.

Firestone, W. A., Mayrowetz, D. & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. Educational Evaluation and Policy Analysis, 20, 95-113.

Fuchs, L. S, Fuchs, D., Karns, K., Hamlett, C. L., & Katzaroff, M. (1999). Mathematics performance assessment in the classroom: Effects on teacher planning and student problem solving. American Educational Research Journal, 36, 609-646.

Geisler, C. (1995). Writing and learning at cross purposes in the academy. In J. Petraglia (Ed.), Reconceiving writing, rethinking writing instruction (pp. 101-120). Mahwah, NJ: Erlbaum.

Gordon, B.M. (Fall 2000). On high stakes testing. AERA Division G Newsletter. [On-line]. Available: aera.net.divisions/g/gnews/falloo-1.htm.

Graves, D. H. (1983). Writing: Teachers & children at work. Exeter, NH: Heinemann.

Koretz, D. & Barron, S. (1998). The validity of gains in scores on the Kentucky instructional results information system (KIRIS). Santa Monica, CA: RAND.

Koretz, D., Mitchell, K., Barron, S., & Keith, S. (1996). Final report: Perceived Effects of the Maryland School Performance Assessment Program. Los Angeles: University of California.

Martin, N. (1992). Language across the curriculum: Where it began and what it promises. In A. Herrington & C. Moran (Eds.), Writing, teaching and learning in the disciplines (pp. 6-21). New York: Modern Language Association.

Masters, J.R. (1992, April). A study of arbitrations in Pennsylvania's writing assessment. (Research Tech Rep. No. 143). Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA. (ERIC Document Reproduction Service No. ED 346 113).

McTighe, J. (1997). What happens between assessments? Educational Leadership 54 (4), 6-12.

Novak, J.R., Herman, J.L. & Gearhart, M. (1996). Establishing validity for performance based assessments: An illustration for collections of student writing. The Journal of Educational Research 89 (4), 220-233.

Paris, S.G. & Windograd, P. (1990). How metacognition can promote academic learning and instruction. In B.F. Jones and L. Idol (Eds.), Dimensions of thinking and cognitive instruction (pp. 15-51). Hillsdale, NJ: Erlbaum.

Pennsylvania System of School Assessment (2000). Writing assessment handbook. Harrisburg, PA: Pennsylvania Department of Education.

Pennsylvania System of School Assessment (1998-1999). Writing assessment handbook. Harrisburg, PA: Pennsylvania Department of Education.

Petraglia, J. (1995). Writing as an unnatural act. In J. Petraglia (Ed.), Reconceiving writing, rethinking instruction (pp. 79-100). Mahwah, NJ: Erlbaum.

Rief, L. (1992). Seeking diversity: Language arts with adolescents. Portsmouth, NH: Heinemann.

Rosenthal, R. & Rosnow, R. L. (1984). Essentials of behavioral research: Methods and data analysis. New York: McGraw-Hill.

Ruth, L. & Murphy, S. (1988). Designing writing tasks for the assessment of writing. Norwood, NJ: Ablex.

Soltis, J.M. & Walberg, H. J. (1989). Thirteen-year-olds' writing achievements: A secondary analysis of the fourth national assessment of writing. Journal of Educational Research 83 (1), 22-29.

Stiggins, R.J. (1997). Student-centered classroom assessment. Upper Saddle River, NJ: Prentice Hall.

Tchudi, S. (1991). Planning and assessing the curriculum in English language arts. Alexandria, VA: Association for Supervision and Curriculum Development.

Torrance, H. (1993). Combining measurement-driven instruction with authentic assessment: Some initial observations of national assessment in England and Wales. Educational Evaluation and Policy Analysis, 15, 81-90.

Vygotsky, L.S. (1962). Thought and language. Cambridge, MA: Harvard University Press.

Walberg, H.J., & Ethington, C.A. (1991). Correlates of writing performance and interest: A U.S. National Assessment Study. Journal of Educational Research 84 (4), 198-203.

White, E. M. (1994). Teaching and assessing writing: Recent advances in understanding, evaluating, and improving student performance. (2nd ed.) San Francisco: Jossey-Bass.

Wiggins, G. (1993). Assessing student performance: exploring the purpose and limits of testing. San Francisco: Jossey-Bass.